Health and Work of the Elderly

Lindeboom, Maarten; Kerkhofs, Marcel

2002

TI 2002-025/3
Tinbergen Institute Discussion Paper

Health and Work of the Elderly

Maarten Lindeboom[1]
Marcel Kerkhofs[2]

[1] Department of Economics, Faculty of Economics and Business Administration, Vrije Universiteit van Amsterdam, and Tinbergen Institute,
[2] OSA Institute for Labour Studies, Tilburg University.

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Health and Work of the Elderly

–

Subjective Health Measures, Reporting Errors and the Endogenous Relationship between Health and Work

Maarten Lindeboom

Free University of Amsterdam and Tinbergen Institute

Marcel Kerkhofs

OSA Institute for Labour Studies

Tilburg University.

This version: February 2002

Address for correspondence: Department of Economics, Free University of Amsterdam, de Boelelaan 1105, 1081 HV The Netherlands, tel (+31-20)-4446033, fax (+31-20)-4446005, email:mlindeboom@econ.vu.nl.
We acknowledge the valuable comments of Rob Alessie, John Rust, Joao Santos Silva, Marcos Vera, Rainer Winkelmann, and participants at the ESSLE 2001 conference in Buch am Ammersee, Germany, the Conference on Social Insurance and Pension Research in Arhus, Denmark, the Applied Micro Economics seminar at University College London and participants of the 2nd European Conference on Health Economics held in Lisbon.

Health and Work of the Elderly

–

Subjective Health Measures, Reporting Errors and the Endogenous Relationship between Health and Work

Abstract

This paper aims to explore the interrelation between health and work decisions of elderly workers, taking the various ways in which health and work can influence each other explicitly into account. For this, two issues are of relevance. Self-assessed health measures are usually at hand in empirical analyses and research indicates that these may be affected by endogenous, state dependent, reporting behaviour. Furthermore, even if an objective health measure is used, it is not likely to be strictly exogenous to labour market status or labour income. Health and labour market variables are correlated because of unobserved individual-specific characteristics (e.g., investments in human capital and health capital). Moreover, one's labour market status is expected to have a (reverse) causal effect on health. A solution to the 'Health and Retirement Nexus' requires an integrated model for work decisions, health production and health reporting mechanisms. We formulate such a model and estimate it on a longitudinal dataset of Dutch elderly.

Jel codes I12, J14
Keywords: Retirement, health, endogeneity, state dependent reporting errors

1 Introduction

This paper aims to explore the interrelation between health and work decisions of elderly workers, taking the various ways in which health and work can influence each other explicitly into account. For this, two issues are of relevance. Firstly, health and work may be endogenously related because there are direct effects of health on work and vice versa, and because there are unobservables that may relate observed health and work outcomes (e.g. through an individual's time preference or previous investments in human capital and health capital). Secondly, in addition to this 'classical' endogenous relationship there are problems with the measurement of the health status in empirical applications. Health, as far as it is related to work is of interest for work decisions and studies in the retirement and the disability application literature mostly have to rely on subjective measures for work related health and more objective indicators of an individual's general health. The problem with objective indicators is that parameter estimates in retirement models are subject to errors in variable bias in the case that these objective measures are not perfectly correlated with work related health. On the other hand, the outcome of a direct question towards an individual's health status may depend on the labour market status of the respondent. There may, for instance, be economic motives or it may be the case that individual's are inclined to give their answer conform to social norms. Reporting health, as a major determinant for inactivity is socially more accepted, and eligibility conditions for some Social Security Benefits, notably Disability Insurance Benefits, are contingent upon bad health. So, individuals out of work may be inclined to overstate health problems. This systematic bias in the reporting behaviour of some individuals implies that it may be dangerous to use subjective health measures to characterise the health condition of the respondents in the sample. It also implies that, used in empirical models of labour supply, these measures tend to lead to an overestimate of the effect of health and an underestimate of the effect of economic incentives[1]. The relevance of this reporting bias will differ across countries and it will depend on the accessibility and generosity of the DI programmes in these countries.

Most European countries have systems that are characterised by strong incentive effects that discourage continued work of elderly employees (see for instance the country studies by Gruber & Wise (1998) and the references listed therein). The Netherlands is an extreme case, both in terms of observed retirement patterns as well as in terms of the characteristics of the institutional setting. Since the mid-seventies Labour force participation rates of elderly males (55 years and older) have dropped about 40% points to a current level of less 40%. Employer provided Early Retirement (ER) schemes

[1] Bound argues on the other hand, that attenuation bias due to measurement error may offset reporting biases and that hence, in practice subjective measures may be preferred over other more objective measures.

allow for retirement at the age of 60, or sometimes even earlier. In addition to these schemes there are Unemployment Insurance schemes (UI) and Disability Insurance schemes (DI). It has been argued that notably the DI system, though not designed for this purpose, has been used explicitly as an alternative route for retirement, with the consent of worker and employer organisation and the DI administrators (see for instance, Aarts & de Jong (1992)).

Strong incentive effects of DI schemes have also been found for a number of other countries. Riphahn (1998) finds evidence for the substitution between UI and DI schemes in Germany. With respect to Disability application behaviour in countries like the United States, Germany and Sweden, it has been argued that labour supply (and labour demand) considerations may have taken part in the decision to apply for benefits. To quote Bound and Burkhauser (1999): “the prevalence of disability transfer recipients per worker has increased at all working ages over the last quarter of the century in the United States and in the Netherlands, Sweden and Germany. This coincides with an increase in both access to and the generosity of publicly provided social insurance and social welfare programs targeted at people with disabilities in the industrialised world.” This implies that in most countries the stock of DI recipients may consist of workers who are in poor health as well as those who are in good health, with consequences for applied research in labour economics and health economics.

A consequence of the above is that an empirical analysis of health and work decisions of elderly workers requires an integrated model for work, health production and health reporting. We formulate such a model and estimate it on Dutch longitudinal data. Estimates of the full model generates consistent estimates of the effect of health and financial incentives on work, the effect of work (history) on health, the size of the reporting errors in subjective health measures and it generates a cleansed work related health measure. It is obvious that these results are of interest for the retirement literature, the health economics literature and public policy. To mention a few issues: The method to purge reporting errors from subjective data is of direct use for the health economics literature on the use of (subjective) quality of life measures. Since the famous Grossman model (1972), various models for health production have been specified. There is however, surprisingly little work that confronts these models with observed data. Most proposed social security and pension system reforms aim at increasing labour force participation rates by making retirement financially less attractive and by suggesting more individualised (capital) funded retirement plans. Seen from that perspective it is for pension and health care policy important to know how increased work efforts will affect current and future health levels.

Our model relates to the models estimated in Sickles and Taubman (1986), Sickles and Yazbeck (1998) and Bound et al (1999). Sickles and Taubman (1986), also have a joint model for health and retirement, but ignore the issue of reporting errors and do not allow work to have an effect on health (production). Sickles and Yazbeck, specify and structurally estimate a model for the demand for leisure and the production of health. The health measure is an index for general health, derived from a range of subjective and objective health measures. Their main interest is in the effect of medical consumption on health and the effect of general health on the demand for leisure by elderly. As argued by Bazzoli (1985) and Bound (1991), health as far as it is related to work is of importance for retirement decisions (the demand for leisure) and estimated effects are subject to errors in variable bias if this measure is not perfectly correlated with work related health. Moreover, they do not deal with the issue of reporting errors[2]. Bound et al. (1999), focuss on the effect of (lagged) health on labour force behaviour of elderly. Subjective health measures are related to a range of health indicators and socio-demographic variables and this model is jointly estimated with a model for labour force transitions. The unobservables of both models are allowed to be freely correlated, but the set of health indicators is assumed to be strictly exogenous for the subjective health measure. This assumption is likely to be violated in practice. Moreover, their model does not allow them to purge the reporting errors from the (subjective) health data and it (therefore) does not help them in getting an unbiased estimate of the effect of health on work.

Our model consists of three stochastically related parts: Firstly, a model for work, where together with the usual socio-economic and demographic variables, financial incentives and health are allowed to affect retirement behaviour, secondly a model for health that relates past labour market outcomes to current health levels, thirdly a model for reporting behaviour that translates the observed subjective health meausure into an health index that is free of reporting errors. In turn this cleansed health index is used in the model for work. The (sub)model for health reporting behaviour is integrated in the full model and it is therefore of direct influence for the outcomes of the model for health and work. Therefore, a seperate section of this paper will be devoted to the construction of this health reporting model and the assumptions underlying it.

There is quite a literature that deals with the issue of systematic response biases in subjective health measures. In order to eliminate the subjective nature of responses to questions about health, various authors have used measures that are believed to be more objective, for instance observed future death of respondents in the sample (Parsons (1982), Anderson and Burkhauser (1985)) or sickness absenteeism records (Burkhauser (1979)). As pointed out previously, health as far as it is associated with work is of importance and parameter estimates in retirement models are subject to errors in variable bias if these objective measures are not perfectly correlated with work related health. The use of lagged

[2] It has to be noted that this is of less relevance for their analysis as their index is expected to be less sensitive to

responses to health questions or an instrumen tal variable method as proposed by Stern (1990) or Aarts & de Jong (1991), Dwyer & Mitchell (1998) are also of little help, since that in itself does not eliminate the state dependent reporting errors. Our health reporting model is an extension of Kerkhofs and Lindeboom (1995) and Kreider (1999).

Kerkhofs & Lindeboom (1995) and Kreider (1999) take a very similar approach. In both studies the group of workers is taken as a benchmark and more objective health measures, such as observed chronic health disorders (Kreider), or a more objective medical test score (Kerkhofs & Lindeboom), are used to filter out the bias relative to the group of workers. The general idea is that workers have no incentives to report with error. The fundamental assumption is that the objective health measure acts as a sufficient statistic for the effect of work on (a latent index for) true health, and that therefore remaining systematic differences between the subjective and objective measures across the labour market states can be attributed to reporting errors. Both approaches allow for different response behaviour of individuals across different labour market states, and therefore differ from studies that use an instrumenting procedure that does not exploit the information from different groups on the labour market explicitly (such as Stern (1990), Aarts & de Jong (1992) and Dwyer & Mitchell (1998).

The main problem with the approaches taken by Kerkhofs & Lindeboom and Kreider is that their approaches will fail to produce correct estimates of the bias in the health responses, in the case that there are unobservables that affect both health and work. The unobservables make included health and labour market variables in the response models in Kerkhofs & Lindeboom effectively endogenous. The same critic applies to Kreider’s paper. He estimates the reporting errors model on workers alone and distillates the reporting errors from a comparison of the results of this (limited information) model with the outcome of a model based on the full sample (i.e. workers and non-workers). In the case that there are unobservables that affect both health and work, differences may reflect differences in reporting behaviour and other behavioural differences that may exist between workers and non-workers.

Benitez-Silva et al (1999) test the hypothesis of Rational Unbiased Reporting of Disability status (labelled as RUR). Their test is based on a comparison of the disability insurance administrators (SSA) award decision (a) with the measure for self-reported disability (d). The idea is that the SSA sets a standard of disability and that this standard becomes common knowledge (a social standard). If individuals exaggerate their health problems, then one would expect the rate of self reported disability to exceed the fraction of those who are ultimately awarded benefits. The hypothesis is tested on a sub sample of individuals who applied for Disability Insurance and for whom both the SSA’s award or deny

reporting errors.

decision is observed. It is concluded that one can not reject the hypothesis that the self-reported disability is an unbiased indicator of the SSA's award decision. We do not pursue this idea for several reasons. First of all, we do not have access to data for individuals for whom we both can observe the DI application decision and the ultimate award decision [3]. Next, it may be questionable whether there exist a common social standard or norm towards DI in countries like the Netherlands where large differences exists in DI rates across different subgroups (i.e. generations) of the population. Finally, and related to the previous point, the stringency of the award/deny decision of the SSA may change over time. This may be particularly relevant in the context of the Netherlands. Easy access to generous DI benefits in the nineteen seventies and nineteen eighties led to a substantive inflow in DI schemes and subsequently to a major DI scheme reform in 1993. Our data are from the period 1993 to 1995, the post 1993 reform period, and it is therefore conceivable that a substantive fraction of the stock of DI recipients would not have qualified according to these new rules [4]. A consequence of this is that the Dutch SSA's (and social) standard at the date of the interview may deviate substantially from the standard applied to the stock of DI recipients, with consequences for reporting behaviour.

In general, it may appear for a large number of OECD countries that standards applied to the formerly enrolled DI recipients may differ from the standards in use at the date of the interview. Disability recipient rates of elderly males for The Netherlands, Germany, Sweden and the U.S. are 24%, 18%, 14% and 7%, respectively. It is unlikely that these differences in observed disability rates can be attributed to cross national differences in the health conditions of populations. These high disability rates for (notably) the European countries have triggered political pressure towards Social Security and Public Pension reforms. [5]

The remainder of the paper starts (section 2) with the introduction of a stylised model to fix ideas about how health and work are allowed to influence each other. This section starts with some

[3] Instead we observe individuals at two points in time. Some of them are observed to be at work, others not UI, DI or ER benefit recipients. This is the typical set-up of most samples used in retirement research.

[4] As a matter of fact, the reform introduced a rule that all DI recipients should be re-examined after 5 years. Those who where already on a DI scheme and who where older than 45 where exempted from this rule as it was believed that a substantive share would have found to be capable of work and for whom it would be difficult to find work.

[5] Though it is not sure whether these arguments may hold for a country like the U.S., the procedure proposed in this paper may also be of relevance for models of health and work applied to U.S. data. The methods developed by Benitez-Silva et al, tests whether the conditional probability of a and d equate. There is no method provided to purge the reporting errors from the health measures in case the RUR hypothesis is rejected. Furthermore, in most questions concerning health, individuals are asked to report their reponses on a Likert scale ranging (mostly) from 1 to 5. Overstating of the health problems may also occur at the margins, that is to say that individuals may be inclined, for example, to report level 5 instead of level 4. In this case the distribution of d, which aggregates the response categories, may remain unaltered even if the respondents report with error. In our

remarks on the Dutch institutions, as this is of direct relevance for the stylised model presented in that section. This section ends with the observation that while trying to estimate the model on observed data one usually has to deal with two problems: the classical endogeneity problems and the problem that we do not observe true (work related) health in our model. Section 3 extends on the latter and formulates a health reporting model that relates observed objective and subjective health measures to the relevant health concept, required for the model. We state conditions that allow us to identify systematic reporting errors in survey data. Section 4 presents the empirical implementation of the full model for work, health and reporting behaviour of the elderly. We discuss estimation of the full model and discuss alternative more simple methods to estimate relevant parts of the model. Section 5 presents simulated maximum likelihood estimates of the full model. Section 6 concludes.

2 A stylised model for health and work of the elderly

Some facts of Dutch Institutions and labour supply patterns are of relevance for the model of health and work of elderly workers. As mentioned in the introduction of this paper, in the Netherlands, early retirement is facilitated by social security provided Unemployment Insurance (UI) and Disability Insurance (DI) programmes and employer provided Early Retirement (ER) schemes. Beyond age 57, these schemes provide in general strong incentives towards retirement. For a more extensive discussion we refer to Kapteyn & De Vos (1998). Kapteyn & De Vos (1998) and Lindeboom (1999) calculated implicit tax rates for ER, UI and DI schemes in the Netherlands [6]. These calculations showed that implicit tax rates for ER schemes are more then 70% and that it is financially most attractive to apply for ER benefits at the very moment that a worker becomes eligible for ER benefits. It is important to note that already at age 55 a significant fraction is observed to be out of work (30%). At this age workers are rarely eligible for ER benefits and therefore the larger part of these non-workers are in either UI (47%) or DI (53%) schemes. Maximum implicit tax rates of UI and DI schemes are about 60% and peak at age 58. It is not possible to collect Retirement pensions, while being on UI or DI. Outflow rates from the stock of non-working individuals are extremely low for Dutch elderly. For elderly UI and DI recipients active search for (re)employ ment is not a requirement for eligibility, and ER recipients actually loose retirement benefits upon re-entering employment. This

approach we explicitly deal with ordered non-dichotonous response variables.

[6] Defined as the ratio of the growth in the present discounted value of the retirement income and the yearly gross wages. See also the NBER project on incentives and retirement around the world by Gruber & Wise (1998).

makes UI, DI and ER effectively absorbing retirement states for elderly workers.

Given the above, one could view retirement as a dynamic process in which the decision to stop or continue working depends on a comparison of retirement options that become available over time. Retirement options are characterised by retirement date (age) and route (ER, DI, UI) and consists of packages of retirement years of leisure and the present discounted value of retirement income streams. Health enters the model because it directly affects relative preferences for income and leisure.

More specifically, suppose that individuals start thinking about retirement at age (age) a=0 The end of the horizon is fixed and taken at a=T. For each labour market state we define $U^k_a=U(Y(a),H(a),l(a))$ as the per period utility flow of being in labour market state k at age a, k= ER,DI,UI. U^k_a depends on income, Y, health, H, and leisure l. At each stage of the life cycle the individual chooses the optimal leisure path and retirement exit route that yields the highest expected lifetime utility. If we assume that no savings take place, that retirement states are absorbing and that, while at work, l is fixed, the optimal date of retirement R and retirement route $k \in$ {ER,DI,UI} follows from maximisation of

$$\mathrm{E}_a \sum_{t=a}^{R-1} \beta^{t-a} U(Y(t),H(t),l) + \sum_{t=R}^{T} \beta^{t-a} U(Y^k(t,R),H(t),0) \tag{1}$$

with respect to k and R, subject to the health production function. While out of work, benefit income of the alternative retirement routes may differ and both the level and the time path of the retirement benefit may depend upon the age when the benefits are collected for the first time. Note that this introduces intertemporal non-separability of the optimisation problem.

Another source of intertemporal non-separability may arise from the health production function. If we allow the current stock of health to be dependent on previous health outcomes and current and past choices concerning work, then one may specify:

$$H(a) = h(H(a-1), H(a-2), ..., H(0), l(a), l(a-1), ..., l(0)) \tag{2}$$

In this model, the worker takes into account his or her present health condition and will recognise the effect of work choices on current and future health [7]. The theoretical structure of the model above will

[7] In this simple model we make no distinction between the consumption of non-medical goods and non-medical

not be imposed explicitly on the empirical model. Instead, a quasi-reduced form approach is taken, where equations (1) and (2) serve as the theoretical basis for the empirical specification of the model and where the specifics of the rules of the Social Security and the Early Retirement schemes are explicitly imposed. Before we turn to the empirical specification, two additional comments need to be made.

Firstly, to allow for observed heterogeneity in retirement and health patterns, observed individual characteristics and unobserved (random) components may enter equation (1) and (2) of the model. These may be included to account for, individual heterogeneity, optimisation errors, and/or uncertainty about future events. To start with the health stock equation, $H(0)$ can be interpreted as the initial stock of health that an individual is endowed with and is usually not observed. In empirical applications, one therefore could replace the initial endowment of health with a set of observed individual characteristics that are likely to be correlated with the endowment, such as socio-demographic variables (X), variables measuring health related behaviour (Z) and unobserved factors (γ). It is reasonable to assume that γ is know to the individual, but not observed to the researcher. Similarly, observed individual characteristics X and unobserved components (ξ^k, k=UI,DI,ER) may enter equation (1). Our interpretation will be that the ξ^k 's include both unobserved individual characteristics as well as uncertainty concerning future stopping dates and routes. Consequently, at each age (a) the individual bases his/her retirement decision on a comparison of the expected returns from continued work and immediate retirement at age[8]. Note, furthermore, that γ and ξ^k are most likely to be correlated, as both may include factors that are typically not observed or are measured imperfectly, such as the individuals time preference, attitude towards risk, and previous investments in human and health capital. This will complicate the empirical analyses considerably. We will return to this issue later.

Secondly, of relevance for the work decision is (true) work related health. Instead we observe subjective responses to questions regarding an individual's (work related) health and more objective indicators of general health. What is required, is a model that relates these observed indicators to the true work related health concept. We do this below.

goods and implicitly abstain from the effect of the latter on the development in health. This is not of relevance for the current application, where no attempt is made to identify the structural parameters of the underlying model. Note furthermore that factors like Z and γ will still enter the empirical specification of the health production model via $H(0)$.

[8] As opposed to the case where ξ^k is completely know to the individual. In this case retirement is a single optimization problem concerning retirement date and route, taken at the start of the decision process.

3 A model for Health Reporting

The reported subjective health measure is denoted by H^S and is the respondent's answer to the question "Does your health limit you in your ability to work?" with response categories (1) not at all, (2) a little, (3) have difficulties to do my job, and (4) cannot work at all. For applications in labour supply and retirement models, this work-related measure is most appropriate as it directly relates to the restrictions an individual perceives in performing his job. Though these health measures are typically observed as discrete indicators, we formulate our model in terms of latent variables assumed to generate the observed indicators. This facilitates the discussion below. We define H^* as the latent variable representing true work related health. Rather then one measure for health one could refer to a set of health measures, representing different aspects of health. For ease of exposition we restrict ourselves to a single measure. The key idea of our approach to identify reporting errors is to compare the subjective health measures to an objective measure of health for different groups of the labour market.

A physician-diagnosed report would be the ideal measure of the respondent's health condition. This diagnosis is, however, usually not available in survey data and we have to rely on other sources of more objective information. With respect to a respondent's general health status a more objective measure may be derived from an extensive questionnaire on various (chronic) health conditions and/or health related impediments in performing a large number of daily activities. One of such questionnaires is the Hopkins Symptoms Checklist (HSCL), a test known to have a high rate of external consistency [9]. See section 4 for a more detailed description of this test. The test yields a score and this score will be used as a more objective measure for general health in the empirical applications of section 5. We denote this more objective measure as H^O. Let's focus first on H^O and H^*.

As documented in the introduction, the basic argument in the literature considering the peculiar relationship between subjective health measures and retirement is that commonly used responses to health questions are subject to roughly two forms of possible biases. First, true health may be related to labour market status S (S=Employed, Unemployed, Disabled or Early Retired). This can be a direct causal relationship, or health and labour market status could be indirectly related through unobservables (denoted as type I endogeneity). Secondly, state dependent reporting behaviour could relate the observed subjective measures to the labour market status S (type II endogeneity). By definition H^O and H^* are not affected by reporting behaviour and therefore the following assumption will facilitate us to deal with type II endogeneity, without the need to be explicit about type I endogeneity problems directly.

[9] With this we mean that the HSCL test has been evaluated in an experimental setting, where diagnoses of

Assumption 1 *The objective health measure (H^O) is a sufficient statistic for the impact of current labour market status (S) on true health (H*) conditional on the set of relevant control variables X. More formally:* pdf ($H^*|\, H^O, S, X$) $\cong$ pdf ($H^*|\, H^O, X$).

This is equivalent with stating that, with respect to type I endogeneity, S affects H^* and H^O (conditional on X) in the same way and hence that pdf ($H^*|\, H^O, X$) is identical for all respondents, irrespective of their value of S. The idea of the reporting model may now become clear. If this assumption holds, then any difference in the subjective measure (H^S) that exists across labour market states (S), after that we have controlled for an objective measure H^O and additional controls (X) must come from reporting errors. In line with this reasoning we may write:

$$H^S = f(H^O, S, X, \varepsilon\,;\, \omega) \qquad (3)$$

where ε is an unobservable and ω a parameter vector. Under the assumption, S captures reporting behaviour. We postpone the empirical implementation of (3) to section 5. It is important that H^O is sufficiently objective. If not, the model will tend to underestimate the true effect of state dependent reporting errors. In case it is sufficiently objective (i.e. its dependence on S does not differ from the dependence of H^* on S) but it is inaccurately measured, then this will be captured by X[10]. In addition, X will also capture any differences that there may exist in reporting behaviour across different socio-demographic groups (for instance between men and women). Identification of the state dependent reporting errors in subjective health variables requires a normalisati on. We believe that as a natural choice the group of employed respondents could be considered, since for this group neither financial incentives nor any social legitimisation to report with error exists [11].

Equation (3) can be used to assess the relative importance of reporting errors in health responses and estimates from this equation could be used to generate cleansed health measures that could be used in additional analyses. It has to be noted, however, that direct estimation of (3) is not straightforward, as

---

physicians where compared with the HSCL test outcomes.

[10] Other more objective indicators could be used, such as observed mortality rates in the panel, whether individuals have certain chronic conditions, or the number of doctors visit in the past 12 months. Though all these measures are clearly more objective then a range of questions to an individual's health status, it is likley that they are too specific to serve as measure of health in our model.

[11] This assumption would be violated in case currently employed workers respond in anticipation to future non-participation.

it is conceivable that ε includes wrongly omitted factors that are correlated with labour market status S and the health measure H^O. In essence, type I endogeneity problems return into the model. So, even if we are only interested in the size of the reporting errors, we still need to add a model for S and H^O.[12]

4 Data

Data are obtained from the first two waves of the CERRA panel survey. The CERRA panel survey is a Dutch survey, specifically designed for the analysis of health and retirement issues. It resembles the Michigan Survey Centre's well known Health and Retirement Survey (HRS). The first wave was fielded in the fall of 1993 and consists of 4727 households in which the head of the household (i.e., the main income earner) was between 43 and 63 years of age at the date of the interview. In each household both head and partner, if present, were interviewed. In the fall of 1995 the same respondents were contacted for a second interview. Approximately 74% of the first wave respondents participated in the second wave, which resulted in about 3500 households. For each wave extensive information is obtained on labour history and current labour market status, sources of income, attitude towards retirement, housing, health and a variety of socio-economic variables.

Internal evaluations of item non-response and representativeness of the first wave of data show them to be of high quality. In general, item non-response was not a problem. Non-response was, however, relatively high for the income questions, with a non-response rate of up to 30 percent for some income sources. The CERRA data were compared to data from the Netherlands Central Bureau of Statistics and found to be comparable based on age, sex, labour market status, and education.

The health variables in the sample contain, among others, commonly used subjective measures such as answers to the questions 'how good would you rate your health' and 'does your health limit you in your ability to work'. Less subjective measures like the number of visits to a physician in the past 12 months, whether one was hospitalised in the past 12 months, whether one has experienced a chronic condition and the outcome of the Hopkins Symptom Checklist (HSCL). The HSCL is a validated objective test of general health used in the medical sciences to assess the psycho-neurotic and somatic pathology of patients (respondents). The HSCL consists of 57 items and is known to have an excellent rate of external consistency, meaning that the test results are highly correlated with objective medical

[12] This does not imply that we need to estimate the complete model, even if we are only interested in the health reporting meachanism. In appendix B we discuss (simple) methods to consistently estimate the health reporting model.

reports on the patients' health status. The responses to these 57 questions result in a mental score, a physical score and a total health score. In our analyses we will use the total health score. The advantage of this HSCL measure over a subjectively, self-assessed health measures is that it is free of (or at least less sensitive to) reporting errors that may depend upon the respondent's labour market status. Tables A1-A3 of the appendix provide descriptive measures of variables used in the analyses, average HSCL scores for different groups on the labour market and average HSCL scores for different responses to the subjective questions regarding health.

5 Empirical implementation of the model and results

5.1 *Empirical Implementation*

The core of the empirical model of this section is based on the stylised model of section 2, though no attempt will be made to fully implement the structural parameters of the theoretical model in the empirical implementation. Given the model structure, the workers retirement problem can be written as a sequence of per period decisions based on a comparison of the value of stop working, $V^k(a) = U^k_a + \beta V^k(a+1)$, $k \in A(a)$, for a given set of available options $A(a) \subseteq \{ER, UI, DI\}$ with the value of continued work, $V^w(a) = U^w_a + \beta E \max\{ \{V^k(a+1)\}_{k \in Aa+1}, V^w(a+1)\}$. The availability of the retirement options is determined by the eligibility rules of the UI, DI and ER schemes. We approximate the value functions by a linear function of true work related health (H^*), a range of socio-demographic variables (X), a function of current and future retirement and work incomes ($f(Y;A)$) and unobservables μ^k and ξ^k. More specifically, for each available option k, $k \in A(a) \cup W$ one may write:

$$V_{it}^k = X_{it}'\varphi_0^k + \varphi_1^k H_{it}^* + \varphi_2^k f(Y_{it}^k; A(a)) + \mu_i^k + \xi_{it}^k \tag{4a}$$

The term μ^k is assumed to be a time constant unobserved component, known to the individual, but not to the researcher. ξ^k is taken as an extreme value distributed idiosyncratic shock. The function $f(.;.)$ exhibits some aspects of the forward looking behaviour, but of course it will not be able to capture all of the relevant aspects of $E \max\{ \{V^k(a+1)\}_{k \in A(a+1)}, V^w(a+1)\}$. It excludes for instance future values of the other state variables of the dynamic programming model and therefore the analysis remains essentially reduced form. The individual chooses the labour market state yielding the highest return. Given the distributional assumptions on ξ^k, the retirement model reduces to a multinomial logit, with endogenous regressors H^* and $f(Y;A)$ and unobservables μ^k, $k \in A(a) \cup W$.

We will use instruments for the financial rewards of work and social security benefits ($f(Y;A)$). These wage and benefit constructs are derived from estimates of a wage and participation model and administrative rules from DI, UI and Early Retirement schemes. We return to these constructs later.

If we follow the assumption made in section 3 that pdf $(H^* | H^O, X)$ is the same, irrespective of an individual's labour market state (S), then reporting errors in H_{it}^S can be identified by a comparison of pdf $(H_{it}^S \mid H^O, X)$ across labour market states. H_{it}^S is observed as the ordered categorical variable with responses ranging from 1 (My health does not prevent me from working at all) to 4 (I can't work at all due to health problems). A natural way to model this is by means of an ordered Probit model, where the index $I(H_{it}^O, X_{it}, \delta_i, \nu_{it}; \omega)$ is determined by H^O, X and unobservables δ_i and ν_{it}. The thresholds, represent response behaviour given the index and are modelled as a function of labour market status [13]. More specifically,

$$H_{it}^S = j \quad \Leftrightarrow \quad c_{j-1}(S_{it}; \omega_{2,j-1}) < I(H_{it}^O, X_{it}, \delta_i, \nu_{it}; \omega_1) \leq c_j(S_{it}; \omega_{2,j}), \quad j=1,\ldots,4 \tag{4b}$$

with S_{it} = Work, UI, DI, ER, and $c_0 = -\infty$ and $c_4 = \infty$. The state dependent reporting thresholds c_j are modelled by means of separate dummies for each level of the threshold and labour market status. Disabled, for instance, are expected to have lower thresholds then employed workers, so that for a given value of the index function they are less likely to respond towards good health (category 1) and more likely to respond towards bad health (category 4). We will take the index function as: $I(H_{it}^O, X_{it}, \delta_i, \nu_{it}; \omega) = g^1(H_{it}^O; \omega_{1,0}) + X_{it}'\omega_{1,1} + \delta_i + \nu_{it}$, where $\nu_{it} \sim N(0,1)$. The function $g^1(.)$ is a (third order) polynomial in H_{it}^O. It is interesting to note that the index function ($g^1(H_{it}^O; \omega_{1,0}) + X_{it}'\omega_{1,1} + \delta_i + \nu_{it}$) can be interpreted as an index for the true work related health H_{it}^*, as this index is not dependent on the reporting behaviour. X_{it} includes the usual socio-demographic variables, like age, gender, education etc, but may include also lagged labour market status and other labour market history variables. Hence, apart from reporting errors, estimates of model (4b) will also give interesting results like the causal effect of work history on health[14]. As discussed in section 3, X will capture both the dissimilarity between H^* and H^O (see assumption 1) and differences in reporting behaviour between different socio-demographic groups (see

[13] See also Kerkhofs and Lindeboom (1995) for health reporting and Dustmann and Van Soest (2001) for language proficiency reporting of immigrants.
[14] Of course, provided that labour market history does not have a direct effect on an individuals reporting behaviour There is actually little reason to expect that there will be an aditional effect of labour market history

equation (3)). We will not try to disentangle these two effects. Consequently, one should interpret the effect of X with caution.

Most important for the estimation of (4a) is that the reporting model (4b) can be used to construct an adjusted or 'cleansed' health measure, from which systematic differences in reporting behaviour are eliminated, in particular the reporting behaviour related to the labour market status of the respondent. In modelling the effect of health on labour supply decisions this proxy of H^* is preferred to H^S and H^O, as it incorporates the work related aspects not present in H^O, but avoids spurious correlation due to state-dependent reporting as in H^S. In the labour supply model we therefore use the cleansed health measure based on estimates of ω_{10}, ω_{11} and ω_2.

The unobserved time constant individual component δ_i in (4b) is likely to be correlated with $g^1(H_{it}^O; \omega_0)$, as both H_{it}^S and H_{it}^O are related measures of an underlying health concept. A direct way to take account of this endogeneity is to extend the model with an equation for H_{it}^O. We model H_{it}^O by the empirical counterpart of the health production function (2) of section 2. In this function observed health at a point in time H_{it}^O is taken as a function of the history of work decisions ($_0\int^t l_{iu}\, \mathrm{d}u$), a set of observed characteristics (X_{it}) and unobservables (γ). We only observe health at two points in time, which prevents us from estimating dynamic panel data model. Therefore in our application γ will encompass elements of the initial stock of health and previous investment in health, other then labour supply decisions.[15] More specifically:

$$H_{it}^O = \alpha_0 + \alpha_1\, {}_0\int^t l_{iu}\, \mathrm{d}u + \alpha_2' X_{it} + \gamma_i + \eta_{it} \tag{4c}$$

where η_{it} is an iid error term that is independent of $_0\int^t l_{iu}\, \mathrm{d}u$, X_{it} and γ_i. In the discussion of the results we will pay particular interest in the effect of $\alpha_1\, {}_0\int^t l_{iu}\, \mathrm{d}u$, the effect of work history on current health outcomes.

The full model (4a)-(4c) is estimated simultaneously and enables one to assess the causal effects of health and financial incentives on work, the effect of work history on general health and work related health and the extent to which subjective health measures are biased [16]. Before we turn to the results,

---

variables on reporting behaviour, given that current labour market status variables are included.

[15] Our survey lacks reliable information on the history of behavioural decisions (other then work) that affect health.

[16] Bound et al (1999) estimate equations (6) and (7). Their specifications differ, as they focus is on the effect of lagged health on work decisions. Furthermore, in their counterpart of (7) a set of subjective health measures is regressed on a range of health indicators and reporting mechanisms are not explicitly dealt with. Sickles and Yazbeck (1998) structurally estimate Euler equations, similar to (6) and (8). In their model for the demand for leisure, H_{it} is replaced by a more objective index for general health. They do not (need to) deal with the issue of reporting errors.

however, we have to be somewhat more explicit about some issues concerning estimation of the full model.

As noted above, equation (4a) reduces to a multinomial logit with unobservables μ^k, $k \in A(a) \cup W$, $A(a) \subseteq$ {ER, UI, DI}. In the choice set *A,* we explicitly take into account whether an individual is eligible for an ER scheme. We normalise with respect to the work status *W*, so three unobservables remain.

There are several ways to impute a cleansed measure of health in the full model. These methods are all based on the notion that only the index function and the thresholds of the workers are used to predict the work related health condition as it would have been reported if the respondent would have had a paid job. The first approach is to generate a cleansed ordered response variable taking on the values 1 – 4 using the index function and the worker thresholds. More specifically, the cleansed health measure takes on the value *j* if $\omega_{2,j-1}^{Work} < g^1(H_{it}^O; \omega_{1,0}) + X_{it}'\omega_{1,1} + \delta_i + \nu_{it} < \omega_{2,j}^{Work}$, with $\omega_{2,0}^{Work} = -\infty$ and $\omega_{2,4}^{Work} = \infty$. Optimisation of the likelihood of the full model is not straightforward, when using constructs like these[17]. As an alternative one could use $g^1(H_{it}; \omega_0) + X_{it}'\omega_{1,1} + \delta_i + \nu_{it}$ as an indicator for the true health. A difficulty with this measure is that it will be difficult to identify $X_{it}'\omega_2$ from the included regressors of the model for work, as the index enters linearly in the model for work. Finally, one could derive cleansed measures based on the probability that an individual reports to be in a specific health status, i.e. $\Pr(\omega_{2,j-1}^{Work} < g^1(H_{it}^O; \omega_{1,0}) + X_{it}'\omega_{1,1} + \delta_i < \omega_{2,j}^{Work})$. We chose this construct.

H_{it}^O in equation (8) is measured as the outcome of the Hopkins Symptoms Checklist (see section 4) and it's value ranges from 0 (very healthy) to 171 (very unhealthy) [18]. We will specify (8) as a linear, random effects, model.

The likelihood contributions consist of probabilities associated with the joint event of observing labour market states, subjective health indicators and observed objective health, for each individual at different points in time. We construct individual histories over the period 1991 to 1995. In 1991, some of the workers are still observed to be active in the labour market, whereas others already have retired completely. Previous labour market status (or rather being at work in the previous period) is included in the model for labour market states and we therefore have to correct for the fact that we are sampling form an ongoing process. We specify a separate logit for being at work in 1991. This multinomial logit includes an unobservable ψ_i that is allowed to be correlated with the other included unobserved

[17] A disadvantage of this method is that convergence of the likelihoodfunction of the full model is not likely to occur when one uses standard algorithms. The likelihood optimises simultaneously the model for work, the health reporting model and the health production model. The construction of the health indicators takes place within the each iteration of the optimization procedure. Consequently any change in the parameters of the health reproting model may discontinuously change the cleansed health indicators, which affects the parameters estimates of the model for work.

individual specific effect of the model (δ_i, γ_i, μ_i^{ER}, μ_i^{DI}, μ_i^{UI}). Next, the likelihood function is specified, conditional on the labour market status in 1991. Estimation of the likelihood function requires integration with respect to the probability distribution of (ψ_i, δ_i, γ_i, μ_i^{ER}, μ_i^{DI}, μ_i^{UI}). No closed form solution exists. We use simulation maximum likelihood methods to numerically integrate the unobservables out of the likelihood function (see for instance Stern (1997)). The likelihood function and the implementation of the simulation method used are reported in Appendix C. Appendix B discusses alternative (simpler) estimation methods that may be used if one only wants to estimate the health reporting model. We now turn to the estimates of the full model (4a)-(4c).

5.2 *Results*

The models for work, health reporting and health production are stochastically related. The model for work and health reporting are also directly related as the latter model is used to construct a cleansed health measure (see also above). Moreover, there exists a direct effect of the HSCL index for general health (equation (4c)) on health reporting behaviour (equation (4b)). The three models are estimated jointly and the construction of the cleansed health measure takes place within each iteration of the optimisation procedure. Below we will present the results of the models. We start with the results of the model for work.

The model for work decisions

Table 1a includes estimates of the model for being in a particular labour market state at time t. We take the workers as the reference group and hence parameter estimates of the Unemployed (UI), the Disabled (DI) and the Early Retirees (ER) are relative to this group. Positive coefficients are associated with lower probabilities of being in a particular state. The first panel of the table refers to control variables at time t; these include age, education, retirement income and (cleansed) health. The cleansed health construct is the probability that an individual reports to be in bad health ("have difficulties to do my job" and "cannot work at all" (response categories 3 and 4)), based on the index function and the value of the workers' thresholds (i.e. $Pr(g^1(H_{it}^O; \omega_{1,0}) + X_{it}'\omega_{1,1} + \delta_i > \omega_{2,2}^{Work})$).

The second panel of the table includes variables for those who where at work in the previous period and hence refers to transition from work to out of work. Of relevance for this is the financial reward from continued work and immediate retirement. The included variables measure different

[18] We however, only observe HSCL scores of at most 136 in our data.

aspects of the financial rewards. A variable is included that relates the retirement income flow of a specific exit route, relative to the stream of income associated with continued work up to the age of (mandatory) retirement. The retirement income streams are based on wage constructs (see Heyma (2001) for details) and specific rules of the social security rules and early retirement schemes (see appendix D for a brief description of Dutch social security and ER schemes).

<insert table 1a about here>

In this second part of the table we also included a variable for the availability of an early retirement scheme within the firm. This variable is included to capture some of the effects of the availability of the very generous ER scheme in the future. We expect this variable to decrease the probability of entering into less generous DI and UI schemes as continued work up to the time that an individual becomes eligible may be rewarding. Indeed the effect of the availability of a ER scheme significantly lowers the probability that in individuals enters into DI and hence DI and ER scheme appear to be options that compete with each other. No such effects are found for UI schemes. It is good to note that both ER and UI schemes are financed from sectoral funds and that it was not uncommon for employers to provide layed off workers with a bonus on top of their UI benefit up to the time that they became eligible for an ER scheme, if available within the firm (see for instance the qualitative study of Bolweg & Dijkstra (1995)). Our results do not reject this idea. [19] The coefficient of the financial incentive variables has the correct sign, but is not significant for either of the exit routes. For UI and DI, this may be due to the relatively low number of observed transition for these groups. Alternatively, there may be little variation in the replacement rate as substantial fraction of new UI and DI recipients qualify for wage related benefits of 70% up to the Mandatory retirement age of 65. [20] Finally, average replacement rates of ER schemes in our sample are above 80% of last earned gross wages and postponement of retirement did not lead to higher replacement rates. Consequently, by far the larger part of the workers collect ER benefits at the very moment that they become eligible. Indeed, simple calculations with the multinomial logit model reveal that that the probability of retirement becomes high and that the relative probability of entering into UI or DI becomes negligible, once an individual becomes eligible for an ER scheme. [21] This explains why additional variation in ER

[19] The funds are formed from employer and worker contributions and until recently there was no experience rating. The results also imply that it will in general be difficult to distinguish the separate incentive effects from ER and UI schemes.

[20] See also section 4. Workers of 571/2 years or older qualify for benefits of 70% of gross last earned wages.

[21] In our likelihood function we adopt the choice set of the multinomial logit depending on the eligibility status of

benefits do not add much to the explanatory power of the model. This also indicates that DI and UI, on the one hand, and ER schemes on the other hand, act as a substitutes.

The cleansed health index is the probabilty that an individual reports to be in bad health (category 3 or 4), using the reporting bounds of the workers (we will turn to these results later). The model estimates suggest for each exit route strong significant effects of health on labour market states. A simple multinomial logit (reported in Table A6 of the appendix) with a bad health indicator (category 3 or 4 of the subjective variable) gives health coefficients (with absolute t-values in parentheses) of –0.826 (3.6), -4.179 (17.25) and -0.511 (2.13) for UI, DI and ER, respectively. It is difficult to compare these coefficients directly to the coefficients of table 1a due to differences in the scales of the health measures. If we rescale the health coefficients in table 1a w.r.t. the average difference in the value of the cleansed health measure between respondents reporting bad health and respondents reporting good health the comparable health coefficients in table 1a would be -2.131, -2.261 and -0.571, for UI, DI and ER, respectively. These can be compared directly with the health coefficients of Table A6. The use of self-reported responses of DI recipients leads to exaggerated effects of health on work. It also seems to downplay the effect of financial incentives for DI recipeints. This can be seen from a direct comparison of the effects of the financial incentive variable in table A6 (1.946) with the corresponding coefficient of Table 1a (5.12). For UI the health effects become more pronounced if we correct for reporting bias and the endogenous nature of health. For ER the health effects are hardly effected. The results of the health reporting model may provide additional information on the reasons for these findings.

Health reporting

Table 1b reports results from the health reporting model. An ordered probit model is used to explain responses to a question concerning work related health. A set of controls and a third order polynomial in the objective health variable (H_{it}^{O} , which is the Hopkins symptons Checklist score) are used to generate a (latent) health index. The actual reporting behaviour depends on the index and the ordered probit bounds, which are allowed to depend on labour market status. This ordered probit model is stochastically related to the model for the work and the health stock equation (to be discussed later), but it is also linked to the model for work via the parameters of table 1b [22].

The subjective health measure is a categorical variable ranging form 1 (no problems at all

---

the respondent.

[22] For each worker in the sample, the worker bounds of panel *iii* in table 1b are used to generate a cleansed measure for health. This is done within the optimization procedure used to estimate the full model.

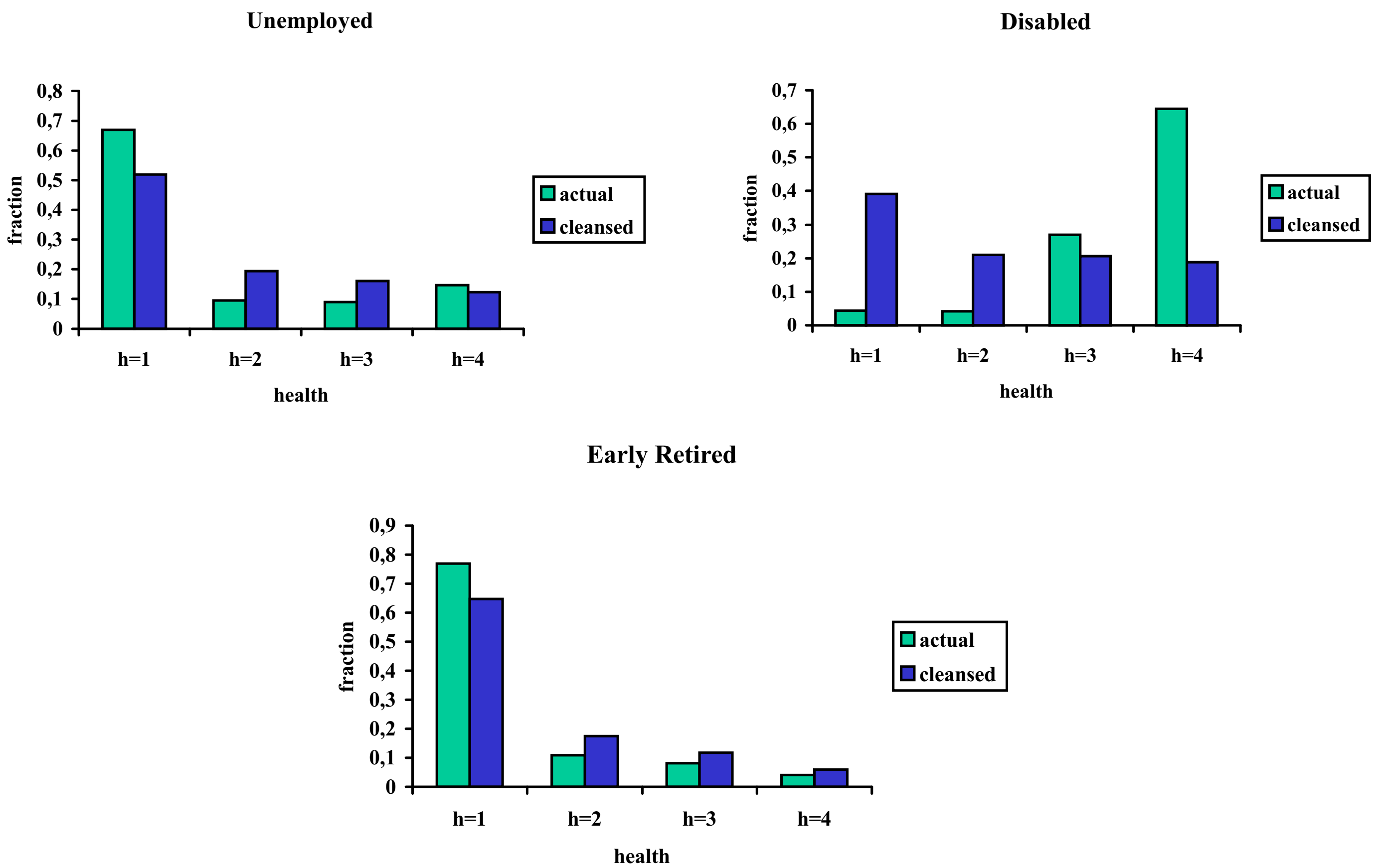
with working) to 4 (can‘t work at all). The lower bound starts at c_1= –0.568 for workers. The other groups on the labour markt have lower bounds relative to the workers, so for disabled for instance the lower bound c_1 equals -1.840 (–0.568 – 1.282). The bounds c_2 and c_3 are taken as an addition to the previous bounds. So for workers, the absolute value of the second bound equals 0.091 (=–0.568 +0.657), for disabled the second bound equals -1.513 (=-1.840+0.327) etc. The results on the bounds indicate that there are indeed strong reporting effects for DI recipients. Conditional on the individuals health status, DI recipients report significantly more towards bad health. It is interesting to note that UI and ER recipients underbounds c_1 are higher then the workers underbound. This means that they are more inclined to report towards good health then the workers. Note however, that for unemployed workers the increments in the bounds are small, implying that unempoyed are less likely to report category 2 ("little problems with work") and 3 ("Difficulties with work"). Relative to the workers, the unemployed are more inclined to report that they can't work at all (the third reporting bound for UI recipients is lower then the third bound for workers). For ER recipients one would say that, apart from response category 1, their reporting behaviour is similar to that of the workers.

<insert table 1b about here>

To make the results more insightful, we used the model to generate predicted responses for a given year and compared the predicitons of the model for different groups on the labour market with actual outcomes (see Appendix E) and with a cleansed measure (see figure 1). The figure in Appendix E indicates that the models describes the data fairly well.

Figure 1 confirms what we have suggested above. Unemployed, report more often to be in very good health (category one) then a worker (given the values of the HSCL score and other exogenous variables), but they are also more likely to report that they can't work at all. The results for the ER are similar to the results for the workers. They report about the same as workers when it comes to the bad health categories (category 3 and 4), but they report to be even more healthy then workers when it comes to the good health category (category 1). Again, it appears that for disabled the biases are large and systematice. Disabled respond heavy towards bad health, and the distribution is almost reversed if one corrects for state-dependent reporting. About 60% of the disabled workers report that they cannot work at all, but when one corrects for the reporting bias this number reduces to about 20%. This result is less extreme as it appears and may even be in line with what one would expect apriori. There is little reason to believe that Dutch elderly are less healthy then their American counter parts[23], but yet about 24% of the Dutch eldery are in a DI scheme. This is more three times as high as the DI rate for the U.S (7%). We estimated a simple ordered probit model, where we do not correct for the fact that the included objective health measure is endogenous. The results of this ordered probit are reported in Table A7. The coefficients on the reporting bounds in this Table indicate that the results become even more extreme. The first bound of the disabled is almost 2 points lower then the first bound of the workers so that calculations with this model would even imply a larger share of healthy DI recipients. This is due to unobserved differences that exists between workers and DI recipients. This also indicates that the previous results of the models estimated by Kerkhofs & Lindeboom (1995) and Kreider (1999), overstate the importance of the reporting bias.

Figure 1 Actual and adjusted ('cleansed') health measures of the health reporting model



[23] Actually, life expectancy of Dutch elderly are among the highest of the world.

As a final remark on the results of the reporting model in Table 1b, one may add that apart from the (polynomial in the) objective health score (HSCL), personal characteristics and an individuals work history are important determinants for the Health index function. Most industrialised countries are currently in a reform of elements of the social security and pension system in order to induce elderly workers to lenghten their work career. In light of this, the effect of work history may be particularly interesting, as this may be informative on the consequences of prolonged work efforts for the average health conditions of future retirees. Work history can affect work related health in two ways. Firstly, directly, these are the coefficients of the polynomials for the number of months ever worked and the number of months worked in the past 10 years. Secondly indirecly, via the effect of the work history variables on the objective health (HSCL) measure. We therefore first discuss the effect for the health production model (the model for the objective (HSCL) health score) and then return to the total (direct and indirect) effect of work on work related health.

The stock of health

Table 1c reports the results for the health production model. Low values of the HSCL score are associated with good health; high values are associated with bad health. The older age groups appear to be the healthiest as are individuals with a partner. It is good to bear in mind that age represents a pure age effect and a cohort effect. The education variables suggest that the reference group (those with only primary school) is the least healthy. Interestingly, and in line with other research, we find that higher incomes are associated with better health.

<insert table 1c about here>

Work history is assumed to be captured by a third order polynomial in the total number of months ever worked and the number of months worked in the past 10 years. The latter is included to capture effects from recent work efforts on health. To illustrate the effect of work history on health production, we calculated the health profile of an average worker with about 15 years of work experience and who remains at work for a number of years. Figure 2a gives a graphical representation.

The values of the calculated HSCL scores of this average worker are in line with observed average HSCL scores of workers in the sample (see table A2). It is clear from the figure that beyond about 300 months (25 years) of work, general health starts to deteriorate with work. The mean of the number of months ever worked is about 350 months and more then 75% of the sample has worked over

300 months, so that effectively for most individuals increased work efforts are associated with worse health outcomes. Figure 2b depicts the same figure for a work related health measure. We used the results of table 1b to calculate the probability that an average individual at work reports to be in bad health ($Pr(g^1(H_{it}^O; \omega_{1,0}) + X_{it}'\omega_{1,1} + \delta_i > \omega_{2,2}^{Work}$). Work experience effects this probability directly and indirectly via the HSCL score (see figure 2a). Figure 2b shows that work related health also starts to deteriorate at about 300 months, and that the effect of work experience may be substantial. The probability that a worker reports to be in bad health more then doubles over the time span considered in the figure.

Figure 2a General Health (HSCL score) and work history (high values are associated with bad health)

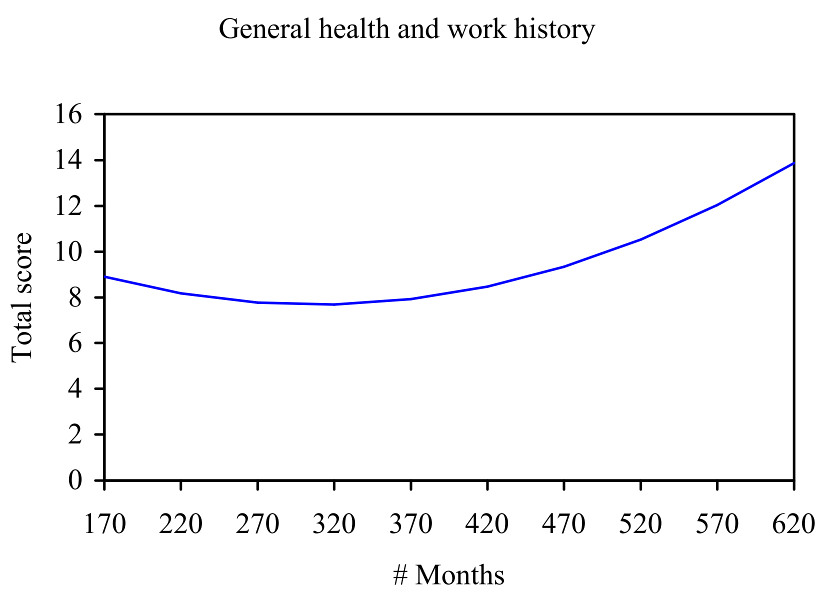
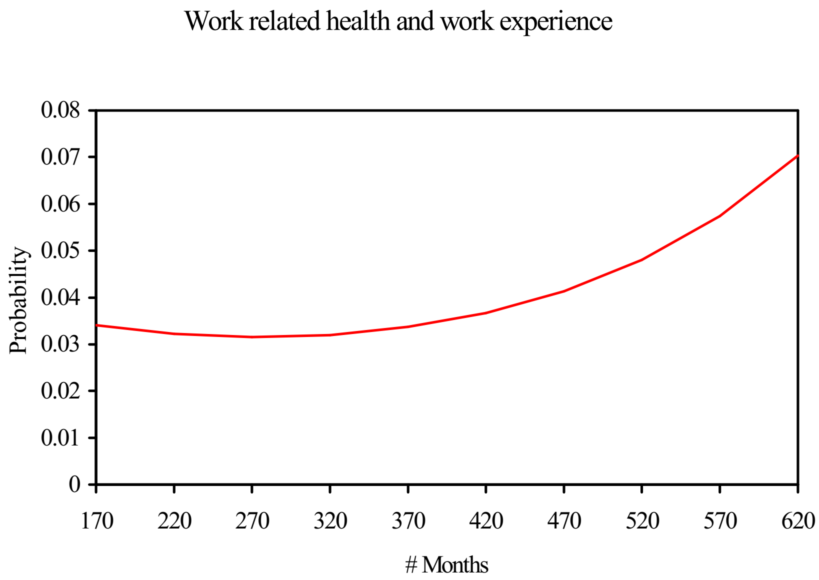


Figure 2b Probability of bad health (work related) for a worker and the effect of work experience



6 Conclusions

This paper aims to explore the interrelation between health and work decisions of elderly workers, taking the various ways in which health and work can influence each other explicitly into account. For this, two issues are of relevance. Self-assessed health measures are usually at hand in empirical analyses and research indicates that these may be affected by endogenous, state dependent, reporting behaviour. Furthermore, even if an objective health measure is used, it is not likely to be strictly exogenous to labour market status or labour income. An analysis of the interrealtion between health and work of elderly therefore requires a model that integrates work decisions, health production and health reporting mechanisms. We formulate such a model. Our model consists of three stochastically related parts: Firstly, a model for work, where together with the usual socio-economic and demographic variables, financial incentives and health are allowed to affect retirement behaviour, secondly a model for health that relates past labour market outcomes to current health levels, thirdly a model for reporting behaviour that translates the observed subjective health meausure into an health index that is free of reporting errors. In turn this cleansed health index is used in the model for work. The model is estimated on a Dutch longitudinal survey using simulated maximum likelihood techniques.

We find that financial incentives are important determinants for the retirement decision. Workers in the Netherlands have strong incentives to take an employer provided Early Retirement (ER) scheme as soon as they become eligible for this scheme. The eligibility for an ER scheme substantially reduce the probability of early outflow through Disability Insurance (DI) and Unemployment Insurance (UI) schemes. Hence it appears that ER on the one hand and DI and UI schemes on the other hand act as substitutes. We furthermore find strong effects of health on (non) work. The use of subjective measures in labour supply models, however, produces biased results. This holds notably for DI recipeints. This finding is in line with the large and systematic biases in self assessed health measures that we find for DI recipients. The health production model reveals that increased work efforts eventually (after about 25 years) lead to a deterioration of health. This finding suggests that pension and social security reforms that aim at increasing labour force participation rates of elderly, may have an adverse effect on the distribution of health among the elderly, with obvious effects for, for instance, health care consumption.

REFERENCES

Aarts L.J.M. & Ph. R. de Jong (1992), *Economic Aspects of disability behaviour*, North Holland, Amsterdam

Anderson, K.H. and R. V. Burkhauser (1985), 'The retirement-health nexus: A new measure of an old puzzle', *Journal of Human Resources*, **20**: 315-330.

Bazzoli, G. (1985), 'The early retirement decision: New empirical evidence on the influence of health', *Journal of Human Resources*, **20**: 214-234.

Benitez-Silva, Hugo, Buschinski, Mosche, Chan, Hiu Man, Cheidvasser, Sofia & John Rust (1999), 'How large is the bias in self-reported disability?', Working paper Brown University.

Bound J (1991) 'Self reported versus objective measures of health in retirement models', *Journal of Human Resources*, **26**, 107-137.

Bound, J., M. Schoenbaum, T. R. Stinebrickner and T. Waidmann (1999), 'The dynamic effects of health on the labor force transitions of older workers', *Labour Economics*, **6**, 179-202.

Burkhauser, R.V. (1979), 'The pension acceptance decision of older workers', *Jounal of Human Resources*, **14**, 63-75.

Dwyer, D. & O. Mitchell (1998), 'Health Problems as Determinants of Retirement: Are self-Rated Measures Endogenous?', *Journal of Health Economics*, **18**, 173-193.

Hsiao, Cheng (1986), *Analysis of Panel Data*, Econometric Society Monographs, Cambridge University Press, Cambridge.

Kapteyn A. & de Vos K. (1999) 'Social Security and Retirement in the Netherlands', in: J. Gruber and D.A. Wise (eds.), *Social Security and Retirement around the World*, NBER, Chicago.

Kerkhofs, M.J.M. & Lindeboom, M (1995) 'Subjective Health measures and State Dependent Reporting Errors', *Health Economics*, **4**, 221-235.

Kerkhofs, M.J.M. and M. Lindeboom (1997), 'Age Related Health Dynamics and Changes in Labour Market Status', *Health Economics*, **6**, 407-424.

Kreider (1999), 'Latent Work Disability and Reporting Bias', *Journal of Human Resources*

Lindeboom, M. & G.J. Van den Berg (1994), 'Heterogeneity in Models for Bivariate Survival: the importance of the mixing distribution', *Journal of the Royal Statistical Society, series B,* **56**, 49-60.

Meuwissen , P.J.J. (1993), 'AOW ontvangers en aanvullend (pensioen) inkomen', (in Dutch) *Statistisch Magazine*, 11-13.

Stern, S. (1989), 'Measuring the effect of disability on labour force participation', *Journal of Human Resources*, **24**, 361-395.

Stern, S. (1997), 'Simulation based estimation', *Journal of Economic Literature,* **35**.

Table 1a Estimates of the full model for work, health and health reporting: Labour market states[1,2].

	Labour market states					
	Unemployed		Disabled		Early Retirement scheme	
	parameter	*t-value*	*parameter*	*t-value*	*parameter*	*t-value*
Constant	10.561	2.70	11.745	2.58	-2.441	-1.56
Age = 56	-1.033	-2.27	-0.144	-0.21	-0.580	-1.48
Age = 57	-0.700	-1.34	0.304	0.40	-1.091	-2.78
Age = 58	-0.270	-0.54	1.006	1.41	-0.520	-1.38
Age = 59	-1.002	-1.93	-0.051	-0.07	-1.681	-4.29
Age = 60	-0.339	-0.59	0.425	0.54	-1.280	-3.21
Age = 61	-0.879	-1.43	-0.158	-0.19	-2.056	-4.34
Age 62 and older	-1.642	-2.82	0.895	1.21	-1.252	-2.64
Education medium	-1.307	-4.13	-0.749	-1.65	-0.667	-2.61
Education higher	-1.270	-3.66	-0.396	-0.83	-0.469	-1.79
Mean (time) Retirement Income	-7.699	-1.77	-8.154	-1.66	1.526	0.98
Health index (H*) [3]	-21.740	-8.51	-29.367	-9.76	-14.649	-6.28
Transition vars (at work at t-1)						
Constant	3.607	5.77	1.729	1.03	2.721	5.37
Part-time work (<32 hrs p.w)	0.365	0.68	-1.016	1.27	1.442	3.74
ln(Σ benefit) – ln(Σ wage) [4]	0.493	0.76	5.120	1.27	0.086	0.62
ER scheme availiable in firm	0.107	0.45	0.748	2.61		
Log likelihood (complete model)			-11157.48			

[1] The reference category is paid work.
[2] The model is estimated with simulated maximum likelihood methods (50 drawings).
[3] Probability of bad health, cleansed measure from the health reporting model.
[4] Stream of benefit incomes up to the age of normal retirement divided by the wage stream associated with continued work (logarithm).

Table 1b Estimates of the full model for work, health and health reporting: Health reporting[1,2].

	Subjective work related health	
	parameter	*t-value*
i Control variables		
Age = 56	0.045	0.55
Age = 57	0.201	2.49
Age = 58	0.186	2.29
Age = 59	0.104	1.22
Age = 60	0.156	1.76
Age = 61	0.040	0.46
Age 62 and older	-0.028	-0.35
Partner	0.040	0.88
White collar worker	-0.125	-3.74
Education medium general	-0.218	-3.73
Education medium vocational	-0.173	-3.27
Education higher general	-0.048	-0.60
Education higher vocational	-0.198	-3.53
Education academic	-0.191	-2.24
Number of months worked in past 10 yrs	0.396	0.99
Number of months worked in past 10 yrs squared	-1.857	-6.48
Number of months ever worked	-1.201	-2.33
Number of months ever worked squared	1.538	3.36
ii Objective Health Measure		
Total score HSCL	1.718	3.82
Total score HSCL squared	-5.357	-3.54
Total score HSCL cubed	3.109	2.36
iii Reporting behaviour		
Work bound c1	-0.568	-4.67
Work bound c2	0.657	18.25
Work bound c3	0.754	11.51
UI bound c1 (addit to Work c1)	0.394	4.05
UI bound c2	0.327	6.77
UI bound c3	0.441	6.84
DI bound c1 (addit to Work c1)	-1.282	-10.45
DI bound c2	0.347	5.28
DI bound c3	1.021	15.40
ER bound c1 (addit to Work c1)	0.339	4.54
ER bound c2	0.453	11.53
ER bound c3	0.704	11.41
Log likelihood (complete model)	-11157.48	

[1] Subjective health is a four category ordered response variable derived from the question “Does your health limit you in the kind and amount of work that you can do?” with answers rangeing from “not at all” to “cannot work”.
[2] The model is estimated with simulated maximum likelihood methods (50 drawings).

Table 1c Estimates of the full model for work, health and health reporting: Health stock equation[1].

	HSCL score[2]	
	parameter	*t-value*
Constant	3.085	33.38
Age = 56	-0.023	-0.29
Age = 57	-0.162	-2.14
Age = 58	-0.118	-1.45
Age = 59	-0.348	-4.18
Age = 60	-0.282	-3.50
Age = 61	-0.465	-5.29
Age 62 and older	-0.491	-6.64
Partner	-0.467	-7.82
Family size	-0.154	-0.67
Education medium general	-0.052	-0.76
Education medium vocational	-0.119	-2.08
Education higher general	-0.036	-0.37
Education higher vocational	-0.023	-0.39
Education academic	0.143	1.49
Income stream	-0.105	-1.82
At work in t-1	-0.052	-0.49
Number of months ever worked	-2.755	-3.36
Number of months ever worked squared	3.126	4.24
Number of months worked in past 10yrs	-0.670	-0.68
Number of months worked in past 10yrs squared	0.086	0.05
Number of months worked in past 10yrs cubed	-0.363	-0.34
s^2 (variance)	1.293	56.23
Log likelihood (complete model)	-11157.48	

[1] The full model is estimated with simulated maximum likelihood methods (50 drawings).
[2] High values are associated with bad health.

Table A1 Sample descriptives.

	Mean	Min	Max
Age in 1993	55.13	42.25	65.00
Education medium general	0.126	0	1
Education medium vocational	0.163	0	1
Education higher general	0.046	0	1
Education higher vocational	0.179	0	1
Education academic	0.059	0	1
White collar worker	0.538	0	1
Family size	2.646	1	10
Number of months ever worked	347	0	707
Early retirement scheme in firm	0.716	0	1
Unemployed (UI)	0.073	0	1
Disabled (DI)	0.176	0	1
On Early Ret. Scheme (ER)	0.195	0	1
Reports bad health[1]	0.140	0	1
Total HSCL score	12.339	0	125
Attrition between wave I & II		0.24	
Number of observations in wave I		3038	

[1] Whether the respondents reports to have difficulties with performing work or that he/she cannot work at all.

Table A2 Mean of total HSCL score & Labour Market State in 1993

Labour market state in 1993	Mean HSCL score
At work	10.01
Unemployed (UI)	11.85
Disabled (DI)	24.81
Early retired (ER)	7.98

Table A3 Mean of total HSCL score & subjective health

Reported health	Mean HSCL score
No problems with work	17.77
Little problems with work	16.43
Difficulties with work	21.55
Cannot work at all	25.00

Table A4 Estimates of the full model for work, health and health reporting: Labour force participation in 1991[1,2].

	parameter	*t-value*
Constant	1.322	2.72
Age =56	-0.825	-4.58
Age = 57	-1.366	-8.06
Age = 58	-1.436	-8.25
Age = 59	-2.293	-11.17
Age 60 and older	-3.047	-12.54
Education medium	0.764	6.09
Education higher	0.848	5.76
Income of continued work	-0.020	-0.13
Eligible for ER scheme	-1.082	-7.88
Log likelihood (complete model)	-11157.48	

[1] Logit model with unobservable for the probability of being at work in 1991.
[2] The full model is estimated with simulated maximum likelihood methods (50 drawings).

Table A5 Estimates of the full model for work, health and health reporting: Parameters of the mixing distribution[1].

	parameter	*t-value*
θ [2]	1.620	14.90
l_{11}	2.428	9.16
l_{12}	0.316	1.06
l_{13}	0.708	3.02
l_{14}	0.361	12.95
l_{15}	-0.592	-5.58
l_{22}	-2.100	-8.28
l_{23}	-0.773	-4.16
l_{24}	-0.344	-13.08
l_{25}	-0.032	-0.30
l_{33}	0.157	0.83
l_{34}	0.202	10.07
l_{35}	0.012	0.10
l_{44}	-0.104	-5.50
l_{45}	0.125	0.88
l_{55}	0.225	1.56
Log likelihood (complete model)	-11157.48	

[1] The full model is estimated with simulated maximum likelihood methods (50 drawings).
[2] θ is the factor of proportionality relating the individual effect in the health stock equation (γ) to the random effect in the health reporting equation (δ), i.e. $\gamma_i = \theta \cdot \delta_i$. The parameters l_{11} to l_{55} are the elements of upper triangular matrix of the Choleski decomposition of the covariance matrix of the random effects. The five remaining random effects (μ^{UI}, μ^{DI}, μ^{ER}, δ, ψ) are associated with the labour market states UI, DI, ER, the health reporting equation and the initial work state (1991) respectively.

Table A6 Estimation results of the multinomial logit model for labour market status[1].

	Labour market states					
	Unemployed		Disabled		Early Retirement scheme	
	parameter	*t-value*	*parameter*	*t-value*	*parameter*	*t-value*
Constant	1.270	0.52	0.548	0.20	-4.868	-4.24
Age = 56	-0.978	-3.51	-0.393	-1.12	-0.622	-1.92
Age = 57	-1.257	-4.38	-0.585	-1.69	-1.461	-4.66
Age = 58	-0.990	-3.11	-0.396	-1.15	-0.938	-3.16
Age = 59	-1.214	-3.61	-0.706	-1.88	-1.890	-5.94
Age = 60	-1.188	-3.27	-0.916	-2.51	-1.741	-5.62
Age = 61	-1.353	-3.53	-0.934	-2.32	-2.243	-6.02
Age 62 and older	-1.863	-4.64	0.164	0.38	-1.368	-3.57
Education medium	-0.224	-1.19	0.116	0.53	-0.255	-1.36
Education higher	-0.138	-0.67	0.542	2.31	0.021	0.11
Mean (time) Retirement income	-3.369	-1.20	-2.215	-0.73	1.682	1.41
Bad health dummie[2]	-0.826	-3.58	-4.179	-17.25	-0.511	-2.13
Transition vars (at work at t-1)						
Constant	4.954	10.89	5.524	6.57	3.649	11.35
Part-time work (<32 hrs p.w)	0.363	0.92	-0.111	-0.29	1.474	4.28
[ln(Σ benefit) – ln(Σ wage)] [3]	0.706	1.32	1.946	0.99	0.101	0.93
ER scheme availiable in firm	0.339	2.07	0.850	4.58		
Log likelihood			-1975.26			

[1] Paid work is the reference category.
[2] If an individual reports to have difficulties to do his or her job or cannot work at all.
[3] Logarithm of the stream of benefit incomes up to the age of normal retirement divided by the stream of wage income associated with continued work.

Table A7 Estimates of the ordered probit model of the subjective work-related health measure[1]

	Subjective work related health	
	parameter	*t-value*
i Control variables		
Age = 56	0.037	0.43
Age = 57	0.173	2.06
Age = 58	0.198	2.39
Age = 59	0.206	2.43
Age = 60	0.229	2.67
Age = 61	0.237	2.76
Age 62 and older	0.178	2.31
Partner	0.198	3.53
White collar worker	-0.150	-3.50
Education medium general	-0.235	-3.79
Education medium vocational	-0.094	-1.80
Education higher general	-0.209	-1.95
Education higher vocational	-0.162	-2.85
Education academic	-0.175	-1.85
Number of months worked in past 10 yrs	-0.887	-1.41
Number of months worked in past 10 yrs squared	0.242	0.53
Number of months ever worked	-0.102	-0.14
Number of months ever worked squared	0.256	0.40
ii Objective Health Measure		
Total score HSCL	7.167	11.94
Total score HSCL squared	-11.902	-5.67
Total score HSCL cubed	6.290	3.52
iii Reporting behaviour		
Work bound c1	0.948	7.90
Work bound c2	0.660	19.06
Work bound c3	0.704	11.71
UI bound c1 (addit to Work c1)	-0.135	-1.36
UI bound c2	0.310	7.04
UI bound c3	0.416	7.21
DI bound c1 (addit to Work c1)	-1.923	-16.70
DI bound c2	0.324	5.66
DI bound c3	0.970	16.36
ER bound c1 (addit to Work c1)	-0.057	-0.69
ER bound c2	0.424	11.73
ER bound c3	0.667	11.70
Log likelihood	-3438.32	

[1] Not controlling for unobserved individual effects and endogeneity of the HSCL variables. Subjective health is a four category ordered response variable based on answers to the question “Does your health limit you in the kind and amount of work that you can do?” that range from “not at all” to “cannot work”.

Appendix B Alternative estimation strategies of the health reporting model

Let's for simplicity assume that we observe a dichotomous variables H_{it}^{S} representing the responses of individual i at time t. We can therefore take a Probit specification for equations (3):

$$H_{it}^{S} = f^{1}(H_{it}^{O}; \omega^{O}) + f^{2}(S_{it}; \omega^{S}) + X_{it}'\omega^{Y} + \delta_i + \varepsilon_{it} \tag{B1}$$

And the health response is defined as $h_{it}^{S}=1$ if $H_{it}^{S} > 0$ and $h_{it}^{S}=0$ otherwise. The specification includes an individual specific components δ_i and it is conceivable that these omitted variables are correlated with included (functions of) H^{O} and S.

The methods presented below do not require specification of the dependence between S and δ_i, instead we only need to specify empirical counterpart of the health production function of section 2 (see also section 4):

$$H_{it}^{O} = \alpha_0 + \alpha_1 \,{}_0\!\int^{t} S_{iu}\, \mathrm{d}u + \alpha_2' X_{it} + \gamma_i + \mu_{it} \tag{B2}$$

where μ_{it} is an iid error term that is independent of ${}_0\!\int^{t} L_{iu}\, \mathrm{d}u$, X_{it} and γ_i. Clearly, when health related decisions and work related decisions are considered simultaneously (as in the model of section 3), then γ_i will be correlated with ${}_0\!\int^{t} S_{iu}\, \mathrm{d}u$.

A naïve approach to estimate (B1) (or (B2)) would be to assume that δ_i^{S} is orthogonal to the included regressors (X_{it}, $f^{1}(H_{it}^{O}; \omega^{O})$, $f^{2}(S_{it}; \omega^{S})$). In this case simple random effect Probit models could be employed to estimate (B1) separately from B2 and an equation for S. In case δ_i^{S} is correlated with either X_{it}, $f^{1}(H_{it}^{O}; \omega^{O})$ or $f^{2}(S_{it}; \omega^{S})$ alternatives need to be considered. One of these alternatives is to use a fixed effect logit specification for (B1). This approach may be appealing as it requires no assumptions on the distribution of the unobservables, nor does it restrict the unobservables to be uncorrelated with the included regressors. A clear drawback is that a large number of observations may be lost in the estimation procedure. The fixed effects are effectively identified on observed changes in individual response behaviour. Our survey only consists of two waves and it is conceivable that of those already out of work at wave 1, only a few would change their response in the next wave.

The correlation between δ_i^{S} and the potentially endogenous variables could also be specified directly, for instance as $\delta_i^{S} = Z_i'\eta + \psi_i$. The vector Z needs to include a set of instruments that capture the correlation between δ_i^{S} and $f^{1}(H_{it}^{O}; \omega^{O})$ and $f^{2}(S_{it}; \omega^{S})$. ψ_i is additional random noise that is independent of Z_i. A straightforward application of Mundlak (1974) would be to take Z as the averages over time of the potential endogenous variables. Especially with relatively short panels as ours, it may be the case that Z is strongly correlated with $f^{1}(H_{it}^{O}; \omega^{O})$ and $f^{2}(S_{it}; \omega^{S})$. The function $f^{2}(S_{it}; \omega^{S})$ represents the effect of state-dependent reporting behaviour and it may be difficult to obtain a precise estimate of this if it is too strongly correlated with Z.

To circumvent this problem we could alternatively, exploit the information that is available in the health stock equation (B2). H_{it}^{O} is measured as the outcome of the HSCL score and ranges from 0 to 171. Therefore (B2) could be estimated using fixed effect regression techniques. In this way dependence of γ_i with the included regressors, of which current labour market status and history are the most prominent variables, is dealt with in the most flexible way. Clearly, the fixed effect γ_i is directly related to S_{it} and H_{it}^{O} and it would therefore serve as a perfect instrument to be included to capture the correlation between δ_i^S and $f^1(H_{it}^{O};\omega^{O})$ and $f^2(S_{it};\omega^{S})$ in equation (B1). So we could specify the dependence as follows: $\delta_i^S = \eta\,\gamma_i + \psi_i$, and substitite this into (B1) and estimate this modified equation equation with standard random effect methods.

More specifically,the fixed effect γ_i can be expressed as $H_{i\bullet}^{O}$ - $\alpha_1 f_{i\bullet}(S)$ - $\alpha_2' X_{i\bullet}$-$\mu_{i\bullet}$. The symbol $f_{i\bullet}(S)$ represents the average over time of ${}_0\!\int^t S_{iu}\,du$, $H_{i\bullet}^{O}$ represents the average over time of $H_{it.}^{O}$, and $X_{i\bullet}$ and $\mu_{i\bullet}$ are defined similarly (see for instance Hsiao (1986)). Essentially, this is the approach suggested by Smith and Blundell (1986).

It is important to note that inclusion of an estimate of the fixed effect introduces additional noise that will certainly be correlated with the labour market status and objective health variables in equation (B1). Therefore, it may be advisable to simultaneously estimate equation (B1) and (B2), thereby acknowledging the expression for γ_i.

Appendix C Likelihood function

The CERRA-survey consists of two waves, held in 1993 and 1991. The 1993 wave contains retrospective information on labour market history. We choose to model labour market states, health production and health reporting in 1993 and 1995. As the model for labour market states is a dynamic one (it includes a variable indicating whether or not the individual was at work in the previous period interacted with financial incentive variables) we in addition have to specify an equation for being at work in 1991.

The conditional logit model for the labour market state includes the unobserved individual components μ_i^{ER}, μ_i^{DI}, μ_i^{UI}. For the probability of having a paid job in the initial year 1991 we include the unobservable ψ_i. The health reporting model and health production models include individual unobserved components δ_i and γ_i, respectively. We assume a one factor error specification for the unobservable of the health production model and the health reporting model, more specifically, $\gamma_i = \theta \cdot \delta_i$. Note that the health production also contains the additional error (see equation (8) in section 5.1) η_{it}. This is convenient for the model as this effectively breaks the direct link of the covariance between the health production and health reporting equation and the variance of the health production equation [24]. Conditional on the unobservables (ψ_i, δ_i, γ_i, μ_i^{ER}, μ_i^{DI}, μ_i^{UI}) the likelihood is simply the product of the likelihood contributions of (1) the multinomial logit model describing whether the respondent's labour market situation in 1993 and 1995, (2) the ordered Probit model of the subjective work-related health measure for 1993 and 1995, (3) the linear regression specification of the HSCL-score in 1993 and 1995 and (4) the logit model describing whether or not the respondent had a paid job in 1991. In order to compute the likelihood we have to take the expected value w.r.t. the random effects (ψ_i, δ_i, μ_i^{ER}, μ_i^{DI}, μ_i^{UI}):

$$L = \int_{\psi} \ldots \int_{\mu^{ER}} \prod_{it} \{ f_1(S_{it} \mid X_{it}, S_{i,t-2}, H_{it}^*, \mu_i^{ER}, \mu_i^{DI}, \mu_i^{UI}) \cdot f_2(H_{it}^S \mid X_{it}, H_{it}^O, S_{it}, \delta_i) \cdot$$

$$f_3(H_{it}^O \mid X_{it}, S_{it}, \theta \cdot \delta_i)\} \cdot f_4(S_{i,1991} \mid X_{i,1991}, \psi_i) \; \mathrm{d}F(\mu_i^{ER}, \mu_i^{DI}, \mu_i^{UI}, \delta_i, \psi_i)$$

In general, no closed form solution exists and we therefore use simulated maximum likelihood methods to numerically integrate the unobservables out of the likelihood function (see for instance Gourieroux & Montfort (1994) and Stern (1997)). We take (ψ_i, δ_i, μ_i^{ER}, μ_i^{DI}, μ_i^{UI}) as jointly normal with expectation zero and covariance matrix Σ=LL', where L is an uppertriangular matrix. In order to evaluate the likelihood L, we first simulate a matrix A of 5 independent standard normal distributed

[24] This would be different for duration models or for discrete choice models, as in these models, with a one factor error specification, the variance of each marginal distributions is restrictively related to the covariance between these two distributions (see for instance Lindeboom & van den Berg (1994)).

normal variables. It follows then that LA are normally distributed with expectation zero and covariance matrix Σ. The parameter θ and the coefficients of L determine the association between the unobservables (ψ_i, δ_i, γ_i, μ_i^{ER}, μ_i^{DI}, μ_i^{UI}) and are estimated along with the other parameters of the model.

Appendix D A brief description to the Dutch system

Dutch benefit programmes can be divided into Social Security benefit programmes and employer provided Early Retirement (ER) programmes. Social Security programmes consists of Unemployment Insurance (UI) and Disability Insurance (DI) programmes. Unemploy ment Insurance programmes can be divided into Unemployment Benefit (UB) program mes, to provide a safety net for those who lose their income due to involuntary unemployment, and social assistance (SA) provisions.

The UB entitlement period depends upon previous job tenure and work experien ce and lasts up to a maximum of 5 years. Benefit replacement rates are a fixed percentage (70%) of previous gross earnings. Benefit recipients have to be in active search for employ ment to maintain (full) benefits. Recipients 57,5 years and older are exempted from the active search requirement. As a result UB is often a source of pre-pension retirement income for elderly workers. At the conclusion of the UB entitlement period, the unemployed can apply for SA. However, the drop in unemploy ment benefit levels may be substantial as SA benefits are seventy percent of minimum wages (the monthly gross minimum wage was 2,163 Dutch guilders in 1994). SA benefits are provided up to the mandatory retirement age (65 years).

Disability Insurance (DI) is provided to protect those who have a physical and/or mental inability to perform gainful employment. Up to the summer of 1993, benefit levels were 70 percent of gross earnings and in practice were provided up to the mandatory retirement age. Though not designed for that purpose, in the past, DI schemes have been used as an exit route for elderly workers (healthy and unhealthy) with consent of the employer, the worker and the DI administrators (see for instance, Aarts & de Jong (1992)). To reduce the number of DI beneficiaries the government tightened DI regulations in the summer of 1993 and introduced a limited benefit entitle ment period and medical examinations at regular times to assess the disability status of the recipient. Due to political pressure beneficiaries 45 years and older were exempted from the tighter rules. Since 1993 the DI entitlement period depends on age and ranges from 0 to 6 years. After this initial entitlement period benefits levels are lowered, according to a function of previous wages, minimum wages and age. [25])). For workers of 58 years and older, full DI benefits are provided up to the mandatory age of retirement (age 65). Despite the efforts to reduce the inflow into DI schemes, the number of DI claimants continued to grow. In 1970 about 200,000 were enrolled in the DI scheme, in 1980 this has grown to 650,000 and continued to grow to about 900,000 now. Since the mid nineteen eighties the economic recovery has led to a growth of the number of jobs and a steady decline in the number of unemployed (currently about 250,000), but over these years the number of DI recipients continued to grow at a constant speed.

Early Retirement (ER) schemes, introduced in the late seventies, are employer provided schemes

[25] Details on the specifics of the UI and DI benefits are available upon request.

and were initially designed as programmes to induce the elderly to retire early in order to make room for young unemployed workers. ER replacement rates vary by sector or even by firm, but are generally financially very attractive. The average replace ment rate is eighty percent of previous gross earnings and in some cases net replacement rates may be close to one. ER eligibility typically depends on age and/or job tenure. Since 1957 all residents of the Netherlands are entitled to a flat rate social security benefit at age 65. The monthly benefit amount is tied to the government-mandated minimum wage. Almost all workers can supplement these basic social security benefits with mandated employer pension benefits.

Appendix E Actual outcomes versus predictions with the health reporting model